

Systematic Gene Search in the Incyte LifeSeq Database

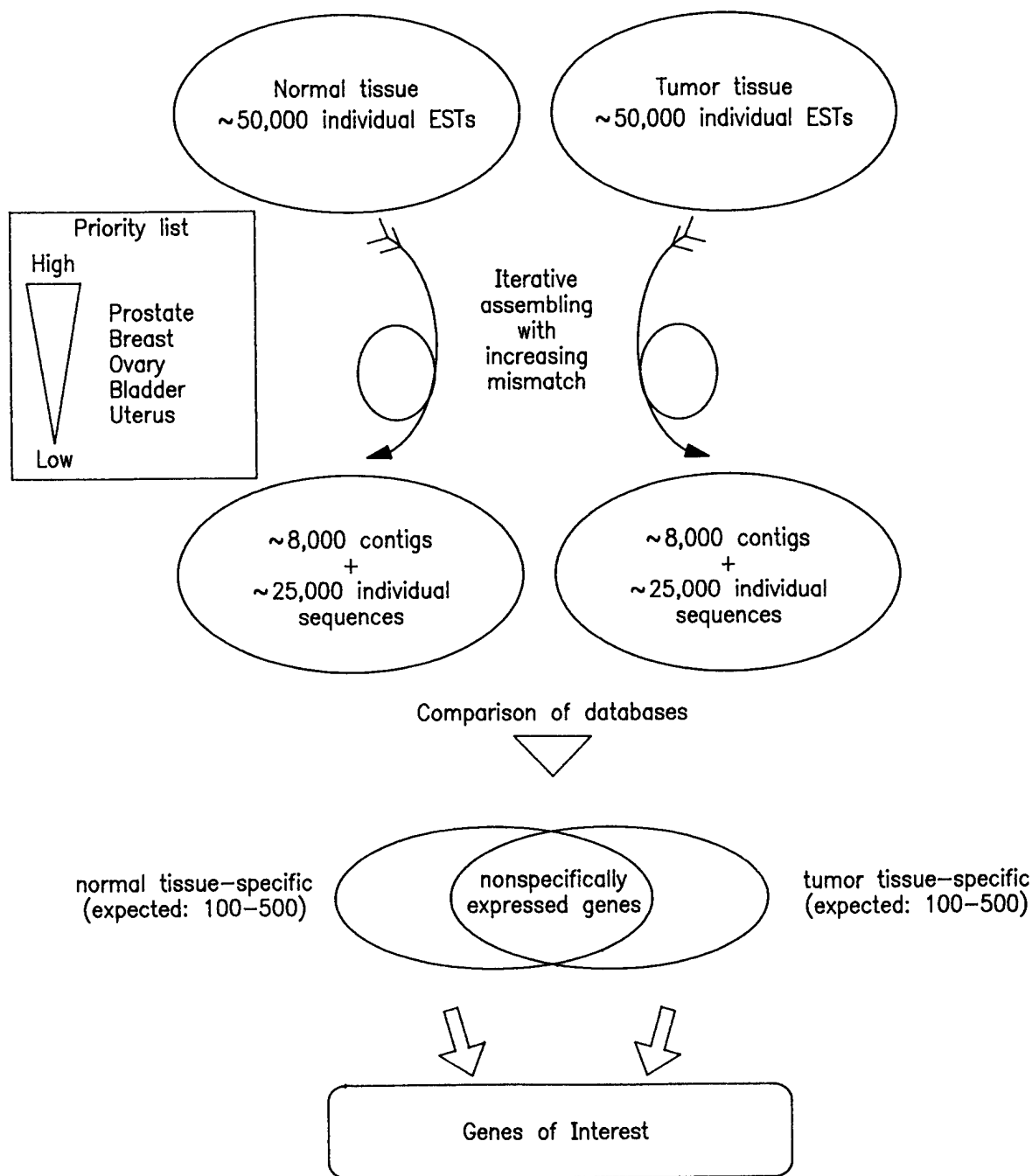


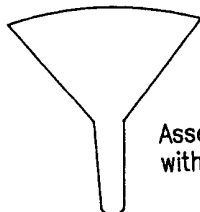
FIG. I

APPROVED	O. G. FIG.	
BY	CLASS	SUBCLASS
DRAFTSMAN		

09/673395

Principle of EST Assembly

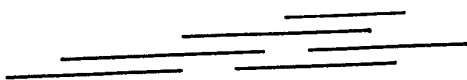
~50,000 ESTs per tissue



Assembly at 0% mismatch
with GAP4 (Staden)



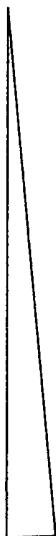
Contigs



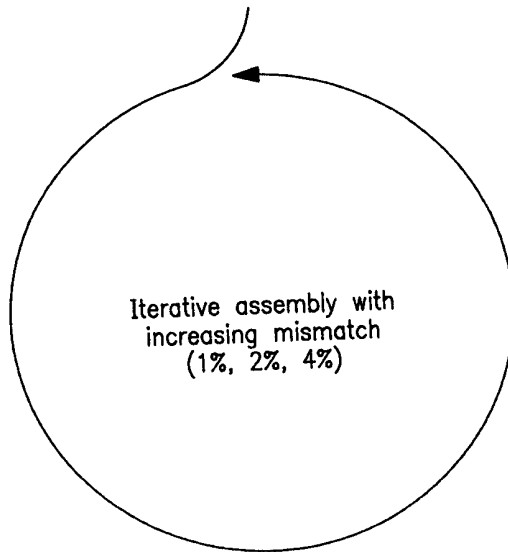
Individual Sequences



Contigs increasing in
number and length



Iterative assembly with
increasing mismatch
(1%, 2%, 4%)



5000-6000 Contigs

~25,000 other individual
sequences



~30,000 consensus-
sequences per tissue

FIG. 2a

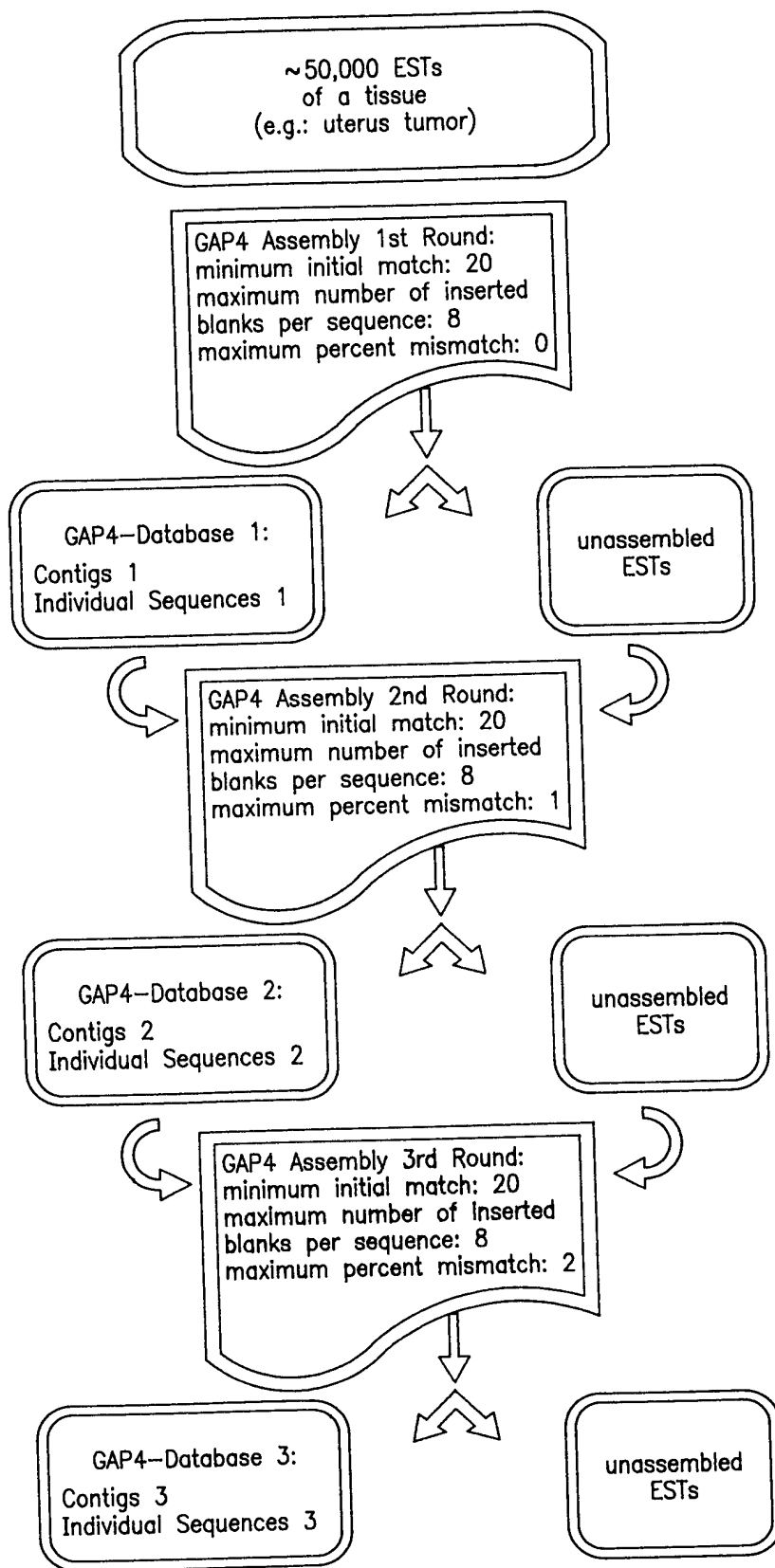


FIG. 2b-1

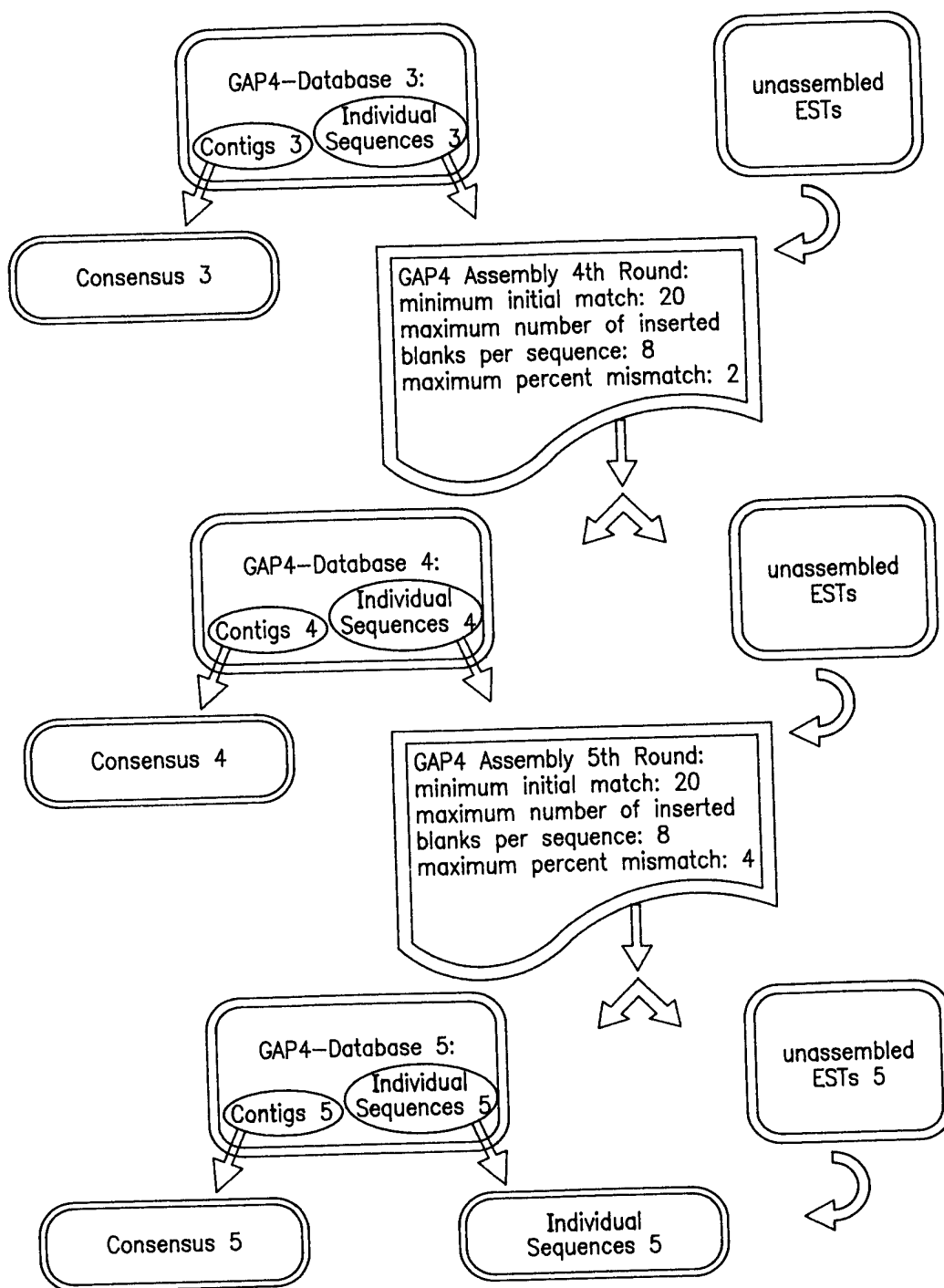


FIG. 2b-2

APPROVED	O. G. FIG.	
BY	CLASS	SUBCLASS
DRAFTSMAN		

09 / 673395

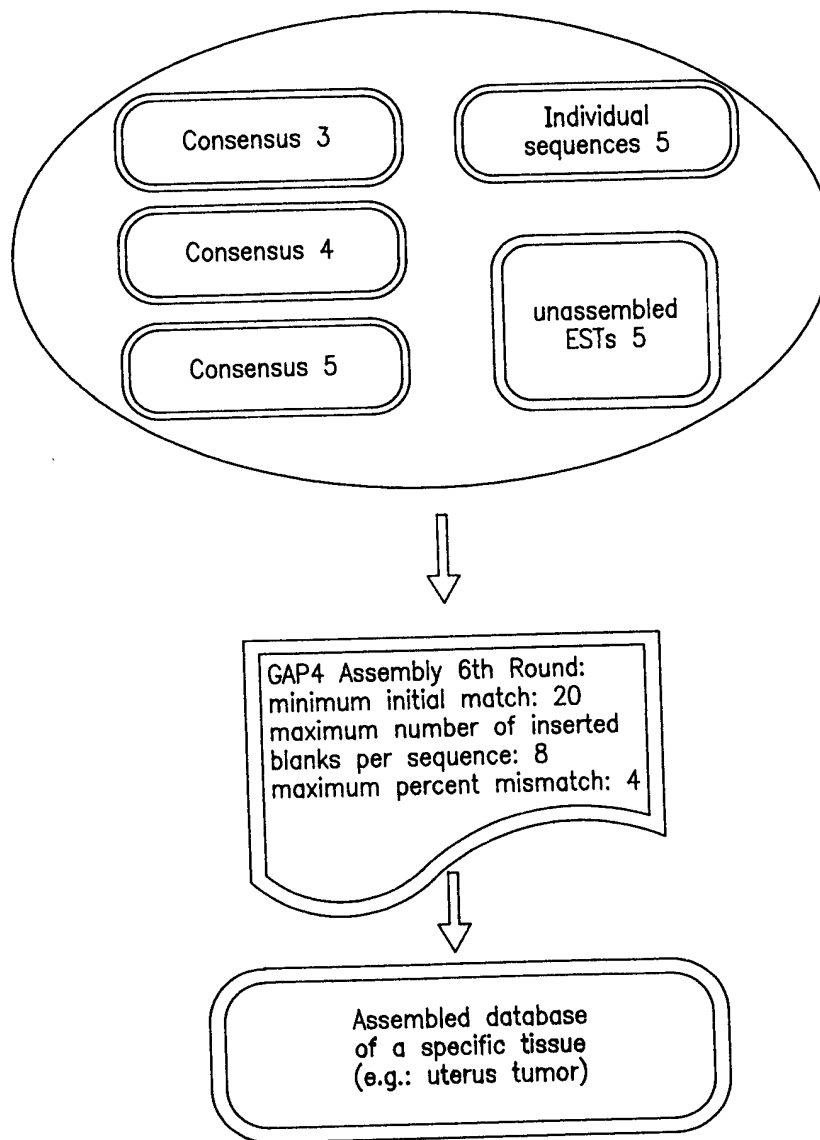


FIG. 2b-3

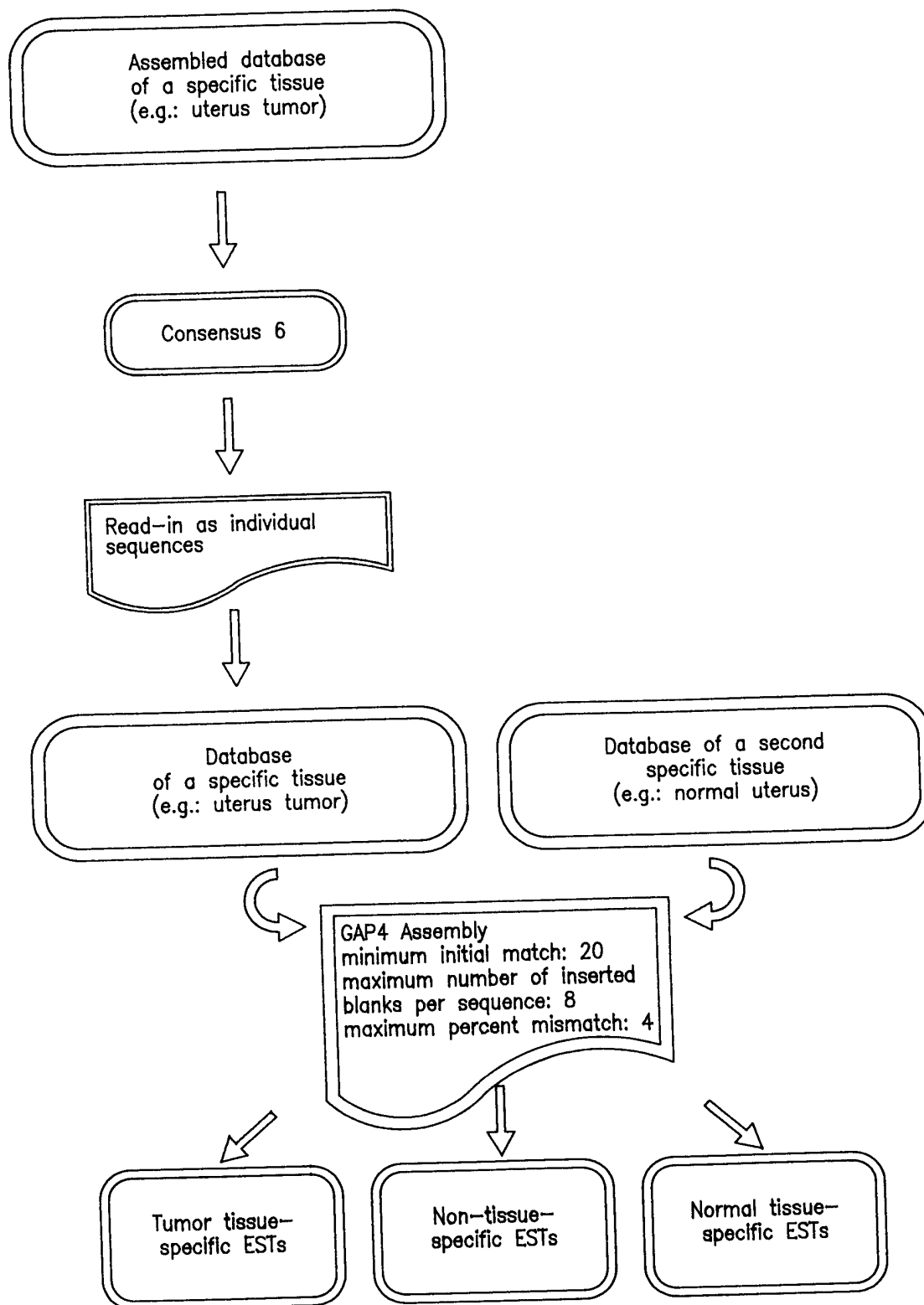


FIG. 2b-4

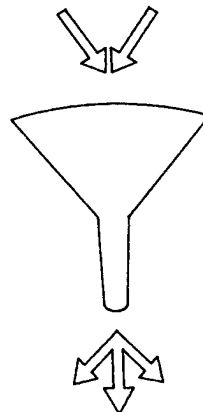
APPROVED	O. G. FIG.	
BY	CLASS	SUBCLASS
DRAFTSMAN		

09 / 673395

In silico subtraction of gene expression in various tissues

~30,000 consensus sequences
normal tissue

~30,000 consensus sequences
tumor tissue



Assembly at 4% mismatch

Normal tissue
Specific genes

Cancer tissue
Specific genes

Genes expressed in both tissues

FIG. 3

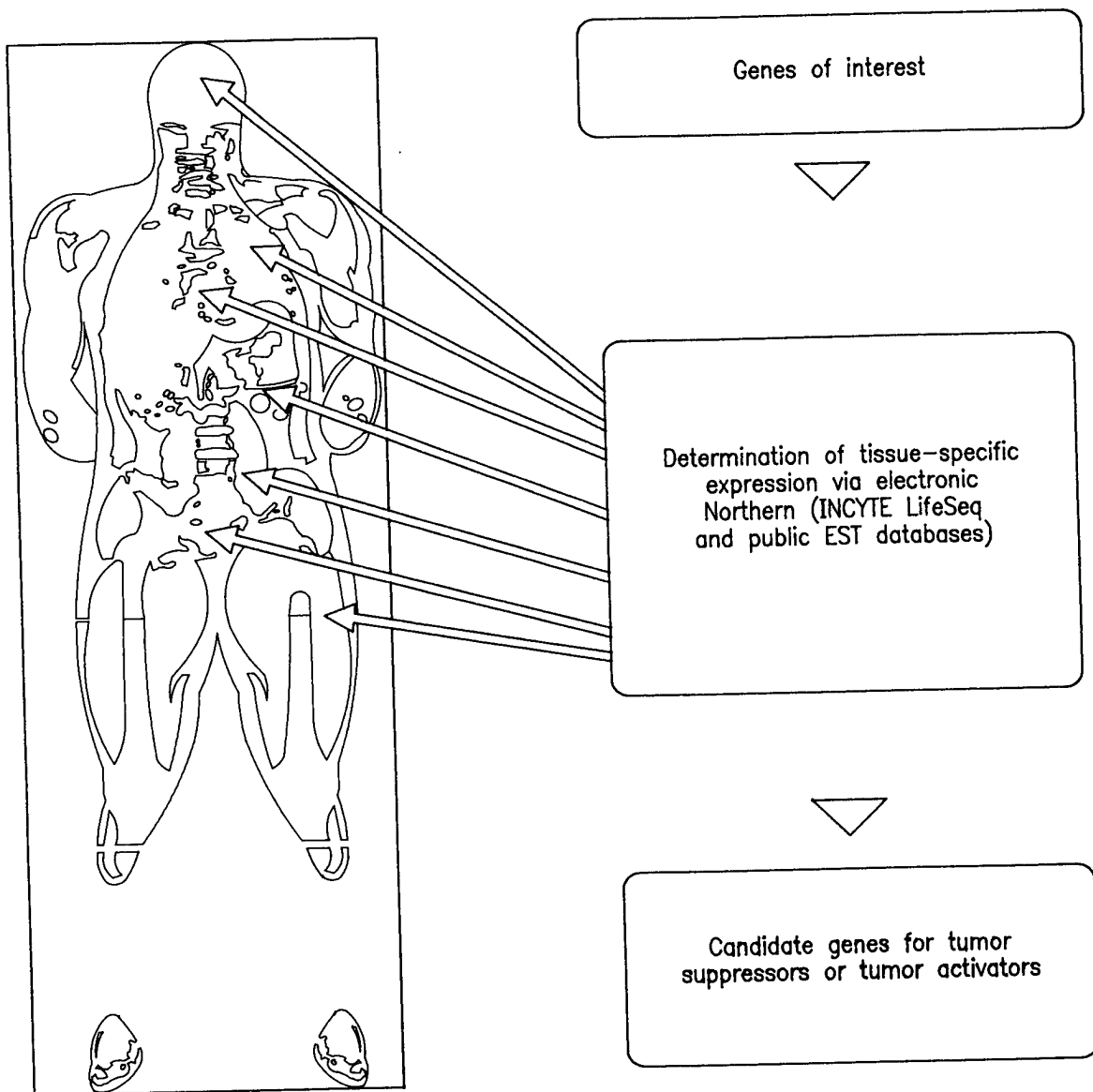


FIG. 4a

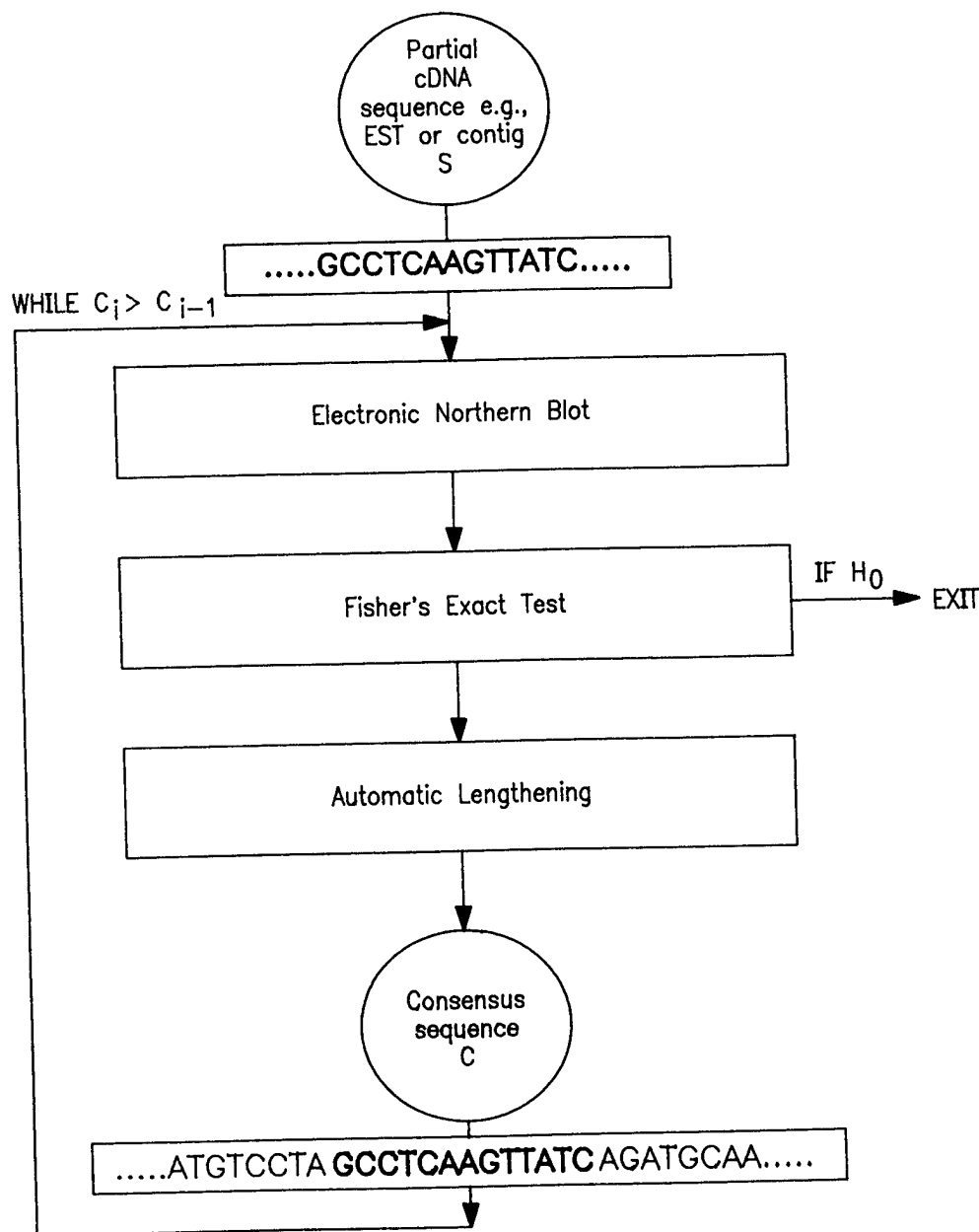


FIG. 4b

09/673395 132700

09673395 122700

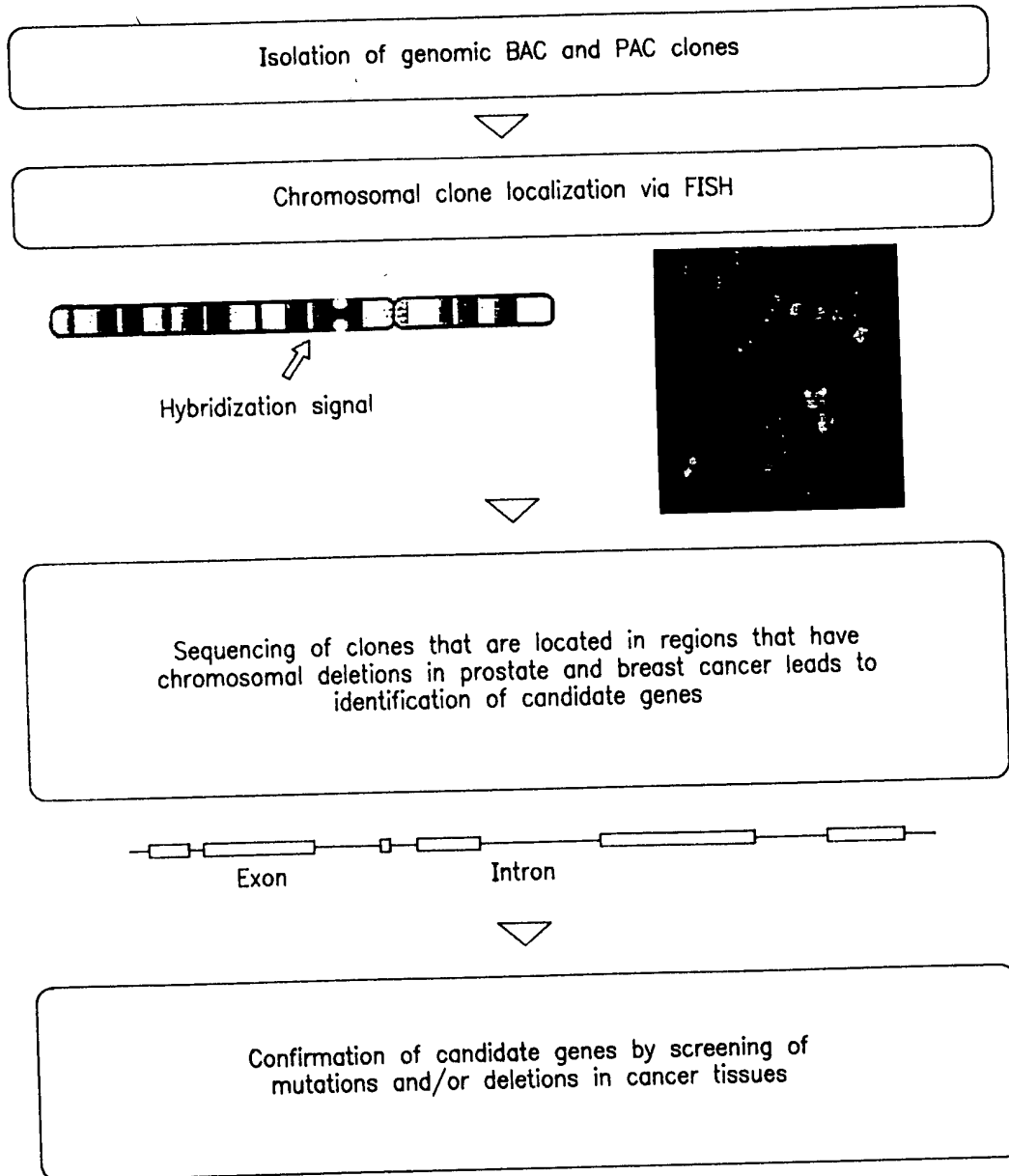


FIG. 5